

Egypt, Arab Rep. - Ti`dad al-Nufus (Population Census) of 1848 - IPUMS Subset

, IPUMS

Report generated on: September 25, 2025

Visit our data catalog at: <https://microdata.worldbank.org/index.php>

Identification

SURVEY ID NUMBER

EGY_1848_PHC_v01_M_v7.5_A_IPUMS

TITLE

Ti`dad al-Nufus (Population Census) of 1848 - IPUMS Subset

ABBREVIATION OR ACRONYM

PHC Egypt 1848 (1%) (IPUMS Harmonized Subset)

COUNTRY/ECONOMY

Name	Country code
Egypt	EGY

STUDY TYPE

Population and Housing Census [hh/popcen] IPUMS International

SERIES INFORMATION

DOI:10.18128/D020.V7.5

KIND OF DATA

Population and Housing Census [hh/popcen]

UNIT OF ANALYSIS

Persons, households, and dwellings

UNITS IDENTIFIED:

- Dwellings: yes
- Vacant Units: no
- Households: yes
- Individuals: yes
- Group quarters: yes

UNIT DESCRIPTIONS:

- Dwellings: Any unit used as a residence, including marginal living quarters not intended as such.
- Households: A person or a group of persons, Egyptians or foreigners, connected or unconnected with blood relation, sharing the same living quarter, food and spending the night together. This includes slaves and free people (including servants). The exception to this are military personnel residing in military barracks, foreign non-Ottoman subjects, and Cairo notables. However, military personnel residing outside barracks are enumerated. Also, foreign non-Ottoman subjects are recorded in the census registers without the household members with a note that they are to be enumerated by their respective consulates. The census order can be found in
- Group quarters: Group quarters, production sites and workplaces, poor shelters.

Version

VERSION DESCRIPTION

Version 7.5. The datasets contain selected variables from the original census microdata plus harmonized variables from the IPUMS-International database.

VERSION DATE

2024-10-05

Scope

NOTES

Additional notes on a sample that is part of this study: Egypt 1848 (1%)

TOPICS

Topic	Vocabulary
Demographic Variables -- PERSON	IPUMS
Nativity and Birthplace Variables -- PERSON	IPUMS
Technical Household Variables -- HOUSEHOLD	IPUMS
Disability Variables -- PERSON	IPUMS
Ethnicity and Language Variables -- PERSON	IPUMS
Geography: A-E Variables -- HOUSEHOLD	IPUMS
Group Quarters Variables -- HOUSEHOLD	IPUMS
Constructed Household Variables -- HOUSEHOLD	IPUMS
Work Variables -- PERSON	IPUMS
Household Economic Variables -- HOUSEHOLD	IPUMS
Technical Person Variables -- PERSON	IPUMS
Geography: Global Variables -- HOUSEHOLD	IPUMS
Dwelling Characteristics Variables -- HOUSEHOLD	IPUMS
Work: Occupation Variables -- PERSON	IPUMS

Coverage

GEOGRAPHIC UNIT

District

UNIVERSE

All individuals (Egyptians and foreigners) who were present within the political boundaries of Egypt. Certain areas are not enumerated. See

Producers and sponsors

PRIMARY INVESTIGATORS

Name	Affiliation
IPUMS	University of Minnesota

Sampling

SAMPLING PROCEDURE

MICRODATA SOURCE:

SAMPLE SIZE (person records): 80028.

SAMPLE DESIGN: Sample is stratified by province. The target sample size in each province is decided based on the pre-determined sampling rate of 8-10 percent in Cairo and Alexandria and 1 percent in all the other provinces. The target province population is based on the 1882 population census. Systematic sampling by page is applied to the entirety of the handwritten registers for each province. Specifically, an initial page within a range of pages (x) is randomly chosen, and the successive pages every "x" pages are selected until the end of the province's registers. The page range (x) is

determined a priori based on the target sample size and the average number of individual records in the page for each province. Person weights are needed to adjust for the different sampling rates across provinces and for the non-enumeration of certain provinces.

WEIGHTING

Sample weights are calculated based on the sample fraction within a specific province.

Data collection

DATES OF DATA COLLECTION

Start	End
1846-01-01	1848-01-01

TIME PERIODS

Start date	End date
1846	1848

DATA COLLECTION MODE

Face-to-face [f2f]

DATA COLLECTION NOTES

de facto, CENSUS DAY: 1846-1848

Questionnaires

QUESTIONNAIRES

Handwritten registers of household and individual characteristics as well as housing conditions.

Access policy

CONTACTS

CONFIDENTIALITY

IPUMS International distributes integrated microdata of individuals and households only by agreement of collaborating national statistical offices and under the strictest of confidence. Before data may be distributed to an individual researcher, an electronic license agreement must be signed and approved. To gain access to the data, a researcher must agree to the following: (1) Implement security measures to prevent unauthorized access to census microdata. Under IPUMS International agreements with collaborating agencies, redistribution of the data to third parties is prohibited. (2) Use the microdata for the exclusive purposes of scholarly research and education. Researchers must explicitly agree to not use microdata acquired for any commercial or income-generating venture. (3) Maintain the confidentiality of persons, households, and other entities. Any attempt to ascertain the identity of persons or households from the microdata is prohibited. Alleging that a person or household has been identified is also prohibited. (4) Report all publications based on these data to IPUMS International, which will in turn pass the information on to the relevant national statistical agencies. Once a project is approved, a password is issued and data may be acquired through the Internet. Penalties for violating the license include: revocation of the license, recall of all microdata acquired, filing of a motion of censure to the appropriate professional organizations, and civil prosecution under the relevant national or international statutes. These safeguards mirror the principles from the Joint ECE/Eurostat Work Session on Statistical Data Confidentiality. Employees of the Minnesota Population Center who work with the census microdata to produce the harmonized database also sign agreements to respect the confidentiality of the data. IPUMS International works with each country's statistical office to minimize the risk of disclosure of respondent information. The details of the confidentiality protections vary across countries, but in all cases, names and detailed geographic information are suppressed and top-codes are imposed on variables such as income that might identify specific persons. In addition, IPUMS International uses a variety of technical procedures to enhance confidentiality protection. These include the following: (1) Swapping an undisclosed fraction of records from one administrative district to another to make positive identification of individuals impossible. (2) Randomizing the placement of households within districts to disguise the order in which individuals were enumerated or the data processed. (3) Aggregating codes of sensitive characteristics (e.g., grouping

together very small ethnic categories) (4) Top- and bottom-coding continuous variables to prevent identification of extreme cases. The safety record for public-use census microdata is apparently perfect. In almost four decades of use, there has not been a single verified breach of statistical confidentiality. The measures implemented by the IPUMS International are designed to extend this record.

ACCESS CONDITIONS

An adapted version of the dataset, harmonized for international comparability, is available from IPUMS International (<https://international.ipums.org/international/>) under the following conditions:

IPUMS International distributes integrated microdata of individuals and households only by agreement of collaborating national statistical offices and under the strictest of confidence. Before data may be distributed to an individual researcher, an electronic license agreement must be signed and approved. To gain access to the data, a researcher must agree to the following:

- (1) Implement security measures to prevent unauthorized access to census microdata. Under IPUMS International agreements with collaborating agencies, redistribution of the data to third parties is prohibited.
- (2) Use the microdata for the exclusive purposes of scholarly research and education. Researchers must explicitly agree to not use microdata acquired for any commercial or income-generating venture.
- (3) Maintain the confidentiality of persons, households, and other entities. Any attempt to ascertain the identity of persons or households from the microdata is prohibited. Alleging that a person or household has been identified is also prohibited.
- (4) Report all publications based on these data to IPUMS International, which will in turn pass the information on to the relevant national statistical agencies.

Once a project is approved, a password is issued and data may be acquired through the Internet. Penalties for violating the license include: revocation of the license, recall of all microdata acquired, filing of a motion of censure to the appropriate professional organizations, and civil prosecution under the relevant national or international statutes.

These safeguards mirror the principles from the Joint ECE/Eurostat Work Session on Statistical Data Confidentiality. Employees of the Minnesota Population Center who work with the census microdata to produce the harmonized database also sign agreements to respect the confidentiality of the data.

CITATION REQUIREMENTS

Digital samples created by Professor Mohamed Saleh (ORCID 0000-0002-2403-9300) from the original Arabic handwritten manuscripts at the National Archives of Egypt. Professor Saleh is the copyright holder of the dataset. For any download and usage of the dataset, please cite: Saleh, Mohamed (2013). A Pre-Colonial Population Brought to Light: Digitization of the Nineteenth Century Egyptian Censuses. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 46(1), 5-18.

ACCESS AUTHORITY

Disclaimer and copyrights

DISCLAIMER

The user of the data acknowledges that the original collector of the data, the authorized distributor of the data, and the relevant funding agency bear no responsibility for use of the data or for interpretations or inferences based upon such uses.

COPYRIGHT

(c) Copyright 1848, Minnesota Population Center

Metadata production

DDI DOCUMENT ID

DDI_EGY_1848_PHC_v01_M_v7.5_A_IPUMS

PRODUCERS

Name	Abbreviation	Affiliation	Role
IPUMS	IPUMS	University of Minnesota	Integration Harmonization Documentation

DATE OF METADATA PRODUCTION

May 21, 2024

DDI DOCUMENT VERSION

Version 7.5 October 2024. NEW FEATURES.

--Historical data from NAPP project now available from IPUMS-International.

--Historical census data from Canada, Denmark, the United Kingdom, Germany, Iceland, Norway, Sweden, and the United States for the period 1703 to 1911 are now available from IPUMS-International. The complete count and sample datasets were previously disseminated by the North Atlantic Population Project (NAPP). Where possible, the data have been integrated into existing IPUMS-International variable coding schema. Some new variables have been created that are available only for these pre-1960 datasets. NAPP data users should note that many NAPP variables are available from IPUMS-International by different names. For a complete list of NAPP variables that have been renamed in IPUMS-International, refer to the crosswalk.

--Individual country shapefiles for the third-level administrative level of geography are now available for a few IPUMS samples.

--New spatially harmonized previous-residence variables at the second administrative level of geography are available for several samples in this data release. More information is available here. Users should note that many older migration variables are available by different names. Refer to this table for a crosswalk of old and corresponding new migration variables.

--IPUMS now hosts the Census Mosaic data collection. Census Mosaic identifies, gathers, harmonizes, and distributes surviving historical census microdata from regions of Continental Europe where complete centralized records are not available. The Mosaic project was founded by a consortium of historical social scientists in Europe. Data can be downloaded as static files from the Census Mosaic website. Although the data are not yet integrated fully into IPUMS International, variables have been standardized and harmonized to be roughly compatible with IPUMS coding structures.

NEW SAMPLES.

--Full-count datasets for Great Britain 1851, 1861, 1871 (Scotland only), 1891, and 1901.

--Full-count dataset for Sweden 1910. Denmark (1845, 1880, and 1885)

--Labor force surveys from Spain and eight new labor force surveys from Italy added to the series.

Newly added countries:

Benin, Cote d'Ivoire, Finland, Guatemala, Honduras, Laos, Lesotho, Mauritius, Myanmar, Papua New Guinea, Russia, Slovak Republic, Suriname, Togo, and Zimbabwe

New samples for:

Bolivia, Cambodia, Chile, Cuba, Cote d'Ivoire, Egypt (1848 and 1868, historical samples), Fiji, Guinea, Ireland, Israel, Italy, Lao PDR, Mexico, Morocco, Nepal, Netherlands, Palestine, Peru, Philippines, Puerto Rico, Rwanda, Senegal, Sierra Leone, South Africa, Switzerland, Uganda, United States, United Kingdom, United States, Vietnam, and Zimbabwe

SUPPLEMENTAL DATA.

Data from censuses from Benin and Lesotho that record individual fertility and/or mortality events were made available in IPUMS-International. These files can be downloaded and linked to data produced by the extract system.

NEW VARIABLES.

--IPUMS-International now provides harmonized and year-specific geography variables for all countries including 13 new samples from Dominican Republic, Germany, Indonesia, Israel, Malaysia, Mongolia, Nicaragua, Nigeria, Palestine, Paraguay, Thailand, United Kingdom, and Uruguay. First-level and second-level year specific geography variables are also available for all countries. IPUMS provides corresponding, downloadable GIS boundary files for all harmonized and year specific geography variables. More information about IPUMS geography variables is available here.

--IPUMS International now provides spatially harmonized previous-residence variables at the first administrative level of geography. The codes for the spatially harmonized previous-residence variables match the spatially harmonized place of current residence. More information is available here.

--IPUMS International provides spatially harmonized previous-residence variables at the first administrative level of geography for all samples; previously available country-specific migration variables at the first administrative level that were

not fully harmonized spatially have been phased out. Spatially harmonized previous-residence variables at the second administrative level of geography are available for selected samples. More information is available [here](#). Users should note that many older migration variables are available by different names. Refer to this table for a crosswalk of old and corresponding new migration variables.

--IPUMS International now provides spatially harmonized previous-residence variables at the first administrative level of geography for all samples. Spatially harmonized previous-residence variables at the second administrative level of geography are available for several samples in this data release. More information is available [here](#). Users should note that many older migration variables are available by different names. Refer to this table for a crosswalk of old and corresponding new migration variables.

--Lower (third) level geography codes and GIS files have been added for Bangladesh, China, Ethiopia, Mali, Rwanda, and Zimbabwe. Some geography codes and labels might have changed for these countries to accommodate the newer lower level geography.

--Added more detailed 3-digit industry and occupation variables for China 2000.

EDITED SAMPLES.

--Revised full-count data for Great Britain 1881

--Revised full-count datasets for Sweden 1890 and 1900. The revision includes the following changes that improve comparability across Sweden datasets:

--Revisions to certain ethnicity and work variables (and the underlying source data): ORIGIN, LABFORCE, OCCHISCO, OCRELATE, OCSTATUS.

--Revisions to unharmonized source variables: SE1890A_HISCOSE, SE1890A_HISCRELSE, SE1890A_HISCSTATSE, SE1890A_OCCMULTISE, SE1900A_HISCOSE, SE1900A_HISCRELSE, SE1900A_HISCSTATSE, SE1900A_OCCMULTISE.

--A new United States 1850 full-count dataset now matches the corresponding dataset distributed by the USA IPUMS data project. The source variable US1850A_0502 (HISTID) provides a linking key to match person records to the USA version of the data. The IPUMS International version of the data contains names, which the USA version cannot distribute.

EDITED VARIABLES.

An error affecting HHWT for South Africa 2007 was corrected. The existing values were adjusted by a factor of 0.01.

AGEMARR was edited to add data for Hungary 1980 and 1990.

Harmonized and year-specific geography variables for Brazil and Colombia have been edited to accommodate for the availability of refined municipal boundaries. Users should be aware that codes and labels have changed in all harmonized and year specific geography variables for these two countries.

Errors affecting BPLSE2 (formerly BLPARSE) for Sweden 1890 and the underlying source variable were corrected. Several thousand cases were incorrectly coded as 258101000. These cases have been updated with the correct code: 258171000.

Harmonized geography variables for Italy, Philippines, Rwanda, and United States have been edited to accommodate new samples. Users should be aware that codes and labels have changed in all harmonized and year specific geography variables for these countries. More information about IPUMS geography variables is available [here](#).

The codes for the source variable RW2002A_0419 were corrected to include 0 and 8 as possible responses, which were previously identified as 'unknown years' within primary education.

Errors affecting EDUCFJ for Fiji 2006 were corrected.

A problem with PERWT for Tanzania 2012 was corrected. The previous weights were adjusted to properly reflect population totals.

MOMLOC, POPLOC, and PARRULE were updated for the United States 2010 and 2015 samples to include additional information on subfamilies. Prior to this correction, persons above age 17 were not receiving links to their co-resident mothers and fathers.

An error affecting codes for the URBAN variable in Egypt 1986 for Cairo, Alexandria, Port-Said, and Suez was corrected.

An error in INCEARN affecting Venezuela 2001 was corrected. Earned income in the source variable VE2001A_0440 is interpreted as a monthly amount, thus adjustments previously applied to convert data from daily or weekly income were suppressed.

All the six Brazil samples in IPUMS International were replaced with higher density samples.

An edited version of the Chile 2017 sample was introduced to correct an error in household breaks.

Errors affecting codes for GEO1_ZA in South Africa 2011 and ENUTS1 in United Kingdom 1991 were corrected.

Harmonized geography variables for Cambodia, Fiji, and Nepal have been edited to accommodate new samples. Users should be aware that codes and labels have changed in all harmonized and year-specific geography variables for these countries. More information about IPUMS geography variables is available [here](#).

An error in PERWT affecting Nepal 2001 was corrected.

Errors affecting a code in GQ for Brazil 2010 and Indonesia 2010 were corrected. Both census samples now identify 1-person units created by splitting a large household.

An error in MARRNUM affecting Indonesia 1976 was corrected. Some codes for GEO1_EG2006 and GEO2_EG2006 were edited.

Harmonized geography variables for Bolivia, Cuba, Guinea, Ireland, Morocco, Palestine, Senegal, South Africa, and Uganda have been edited to accommodate new samples. Users should be aware that codes and labels have changed in all harmonized and year-specific geography variables for these countries. More information about IPUMS geography variables is available [here](#).

An error in INCEARN affecting Brazil 1980 was corrected.

An error in EDATTAIN affecting Ireland 1971 and 1981 was corrected.

A small proportion of person records in Mexico 1960 were re-classified in MIGRATEP based on information about their current and previous residence. These were previously coded to 'different major administrative unit', even though their place of residence suggests that their last move was within the same major administrative unit.

The second-level technician (higher) degrees for Spain 1991, 2001, and 2011 were re-classified into post-secondary technical education in EDATTAIN.

An error affecting codes for SEX for Egypt 1848 and 1868 was corrected. The values for male and female had been reversed.

A problem with HHWT and PERWT for Canada 2011 was corrected. The previous weights were adjusted to properly reflect population totals.

Harmonized geography variables for Cambodia, Lao PDR, Mexico, Peru, Switzerland, Vietnam, Puerto Rico, United Kingdom, and United States have been edited to accommodate new samples. Users should be aware that codes and labels have changed in all harmonized and year-specific geography variables for these countries. More information about IPUMS geography variables is available [here](#).

Harmonized geography variables for Chile and Sierra Leone have been edited to accommodate new samples. Users should be aware that codes and labels have changed in all harmonized and year-specific geography variables for these countries. More information about IPUMS geography variables is available [here](#).

An error affecting codes for COMPUTER for Senegal 2013 was corrected.

An error affecting labels available in IND for Peru 1993 was corrected.

An error affecting codes for persons previously residing abroad for MIG1_5_BO in Bolivia 2001 and 2012 was corrected.

EDUCAR, EDATTAIN, and YRSCHOOL were adjusted in the Argentina samples to incorporate information on completion of education levels in the data harmonization.

HHWT and PERWT were calibrated in Kenya 1979 to properly reflect the population distribution by province.

In GQ (group quarters status), persons residing in hospitals of all types were reclassified to 'institutional group quarters' from 'other group quarters,' making their treatment consistent with GQTYPE.

Errors affecting codes for BPLBJ2 in Benin 1979, 1992, and 2002 were corrected.

Errors affecting codes for GEO2_BR1970 in Brazil 1970 were corrected.

Data Dictionary

Data file	Cases	Variables
EGY1848_PHC-H-H Household records	16	28
EGY1848_PHC-P-H Person records	80028	32

Data file: EGY1848_PHC-H-H

Household records

Cases: 16

Variables: 28

Variables

ID	Name	Label	Question
RECTYPE	RECTYPE	Record type	
COUNTRY	COUNTRY	Country	
YEAR	YEAR	Year	
SAMPLE	SAMPLE	IPUMS sample identifier	
SERIAL	SERIAL	Household serial number	
PERSONS	PERSONS	Number of person records in the household	
HHWT	HHWT	Household weight	
GQ	GQ	Group quarters (collective dwelling) status	
UNREL	UNREL	Number of unrelated persons	
REGIONW	REGIONW	Continent and region of country	
GEO1_EG1848_1868	GEO1_EG1848_1868	Egypt, Province 1848 - 1868 [Level 1]	
GEO1_EG1848	GEO1_EG1848	Egypt, Province 1848 [Level 1]	
GEO2_EG1848_1868	GEO2_EG1848_1868	Egypt, District 1848 - 1868 [Level 2]	
GEO2_EG1848	GEO2_EG1848	Egypt, District 1848 [Level 2]	
OWNERSHIP	OWNERSHIP	Ownership of dwelling [general version]	
OWNERSHIPD	OWNERSHIPD	Ownership of dwelling [detailed version]	
HEADLOC	HEADLOC	Head's location in household	
EG1848A_DWNUM	EG1848A_DWNUM	Dwelling number	
EG1848A_HHNUM	EG1848A_HHNUM	Household number (within dwelling)	
EG1848A_HHN	EG1848A_HHN	Number of households in dwelling	
EG1848A_PERND	EG1848A_PERND	Number of persons in dwelling	
EG1848A_PERN	EG1848A_PERN	Number of persons in household	
EG1848A_YEAR	EG1848A_YEAR	Year	Household variables Year ____ Census year is 1848, although there were data collection in the years 1845, 1846 and 1847. But most of the data was collected in 1848.

ID	Name	Label	Question
EG1848A_REGISCODE	EG1848A_REGISCODE	Register code	Household variables Register code _____ It is the code of the source census register from which the individual record has been digitized. It is the code used in the bookkeeping of the physical census registers at the National Archives of Egypt. Census registers are organized by province, and within each province, by district (qism) in Cairo and Alexandria, quarter (shiyakha) in other urban provinces, and village (nahiya) in rural provinces.
EG1848A_PGNUM	EG1848A_PGNUM	Register page number	Household variables Page number _____ The page number within the original census register.
EG1848A_PROV	EG1848A_PROV	Province	Household variables Province __ The governorate (muhafaza) in urban Egypt or the province (mudiriya) in rural Egypt. It consists of two digits. The province coding follows the administrative division of the 1882 census.
EG1848A_DIST	EG1848A_DIST	District	Household variables District __ The qism in urban provinces and the markaz in rural provinces. It consists of four digits, the first two digits denote the province and the second two denote the district. The district coding follows the administrative division of the 1882 census.
EG1848A_DWTYPE	EG1848A_DWTYPE	Dwelling type (detailed)	Household variables Dwelling type __ Type of dwelling in which the respondent is living in.
EG1848A_DWTYPEGR	EG1848A_DWTYPEGR	Dwelling type (grouped)	Household variables Dwelling type __ Type of dwelling in which the respondent is living in.
EG1848A_DWLEGAL	EG1848A_DWLEGAL	Dwelling legal status (detailed)	Household variables Dwelling legal status Legal status of the ownership of the dwelling. "Owned" could be by a private individual that is not a household member. The endowment or Waqf refers to a donation without transfer of ownership. All dwellings in rural provinces are recorded as "owned"; this variable exhibits more variation in urban provinces. <input type="checkbox"/> Owned <input type="checkbox"/> Endowed Waqf <input type="checkbox"/> Private endowment <input type="checkbox"/> Charity endowment <input type="checkbox"/> Owned endowment <input type="checkbox"/> Mosque endowment <input type="checkbox"/> State ownership <input type="checkbox"/> Guardianship <input type="checkbox"/> Leased <input type="checkbox"/> Warranted

ID	Name	Label	Question
EG1848A_DWLEGALGR	EG1848A_DWLEGALGR	Dwelling legal status (grouped)	Household variables Dwelling legal status Legal status of the ownership of the dwelling. "Owned" could be by a private individual that is not a household member. The endowment or Waqf refers to a donation without transfer of ownership. All dwellings in rural provinces are recorded as "owned"; this variable exhibits more variation in urban provinces. <input type="checkbox"/> Owned <input type="checkbox"/> Endowed Waqf <input type="checkbox"/> Private endowment <input type="checkbox"/> Charity endowment <input type="checkbox"/> Owned endowment <input type="checkbox"/> Mosque endowment <input type="checkbox"/> State ownership <input type="checkbox"/> Guardianship <input type="checkbox"/> Leased <input type="checkbox"/> Warranted
EG1848A_HHNO	EG1848A_HHNO	Number of households within dwelling	Household variables Number of households within dwelling _ _ _ The number of households living within a dwelling
EG1848A_PERNO	EG1848A_PERNO	Number of persons within dwelling	Household variables Number of persons within dwelling _ _ _ _ Number of individuals living in the dwelling.

Total: 33

Data file: EGY1848_PHC-P-H

Person records

Cases: 80028

Variables: 32

Variables

ID	Name	Label	Question
PERNUM	PERNUM	Person number	
PERWT	PERWT	Person weight	
RELATE	RELATE	Relationship to household head [general version]	
RELATED	RELATED	Relationship to household head [detailed version]	
RELATEH	RELATEH	Relationship to household head, historical	
AGE	AGE	Age	
AGE2	AGE2	Age, grouped into intervals	
SEX	SEX	Sex	
CITIZEN	CITIZEN	Citizenship	
BPLEG2	BPLEG2	Province of birth, Egypt (historical)	
BPLEG4	BPLEG4	District of birth, Egypt (historical)	
ETHNICEG	ETHNICEG	Ethnicity, Egypt (historical)	
LABFORCE	LABFORCE	Labor force participation	
OCCHISCO	OCCHISCO	Occupation, HISCO classification	
DISABLED	DISABLED	Disability status	
EG1848A_PERNUM	EG1848A_PERNUM	Person number (within household)	
EG1848A_AGE	EG1848A_AGE	Age	Person variables Age ___ Respondent's age. Women in Alexandria and Cairo reported only "juvenile" or "adult". The original text included information in years and months for children, which was rounded up here to a whole year if they reported six or more months.
EG1848A_AGEDAYS	EG1848A_AGEDAYS	Age in days	Person variables Age in days __ Children's age in days, for respondents 2 years old or younger.

ID	Name	Label	Question
EG1848A_LEGALSTATUS	EG1848A_LEGALSTATUS	Legal status	<p>Person variables Legal status The legal status of the respondent.</p> <p><input type="checkbox"/> Free <input type="checkbox"/> Slave <input type="checkbox"/> Emancipated slave/Mamluk</p>
EG1848A_NATION	EG1848A_NATION	Nationality	<p>Person variables Nationality The nationality of the respondent. "Under government control" in the original text means Egyptian.</p> <p><input type="checkbox"/> Egyptian <input type="checkbox"/> Non-Egyptian</p>
EG1848A_OCCHISCO	EG1848A_OCCHISCO	Occupation, HISCO (5-digits)	<p>Person variables Occupation _ _ _ _ _ Respondent's occupation coded using the Historical International Standard Classification of Occupations (HISCO). The first digit represents major groups classification. Minor groups classification corresponds to the first two digits of the code. Occupations specific to the Egyptian context that do not have an equivalent in HISCO were assigned special codes under a given HISCO major or minor group.</p>
EG1848A_OCCHISCO2	EG1848A_OCCHISCO2	Occupation, HISCO minor groups	<p>Person variables Occupation _ _ _ _ _ Respondent's occupation coded using the Historical International Standard Classification of Occupations (HISCO). The first digit represents major groups classification. Minor groups classification corresponds to the first two digits of the code. Occupations specific to the Egyptian context that do not have an equivalent in HISCO were assigned special codes under a given HISCO major or minor group.</p>
EG1848A_OCCHISCO3	EG1848A_OCCHISCO3	Occupation, HISCO major groups	<p>Person variables Occupation _ _ _ _ _ Respondent's occupation coded using the Historical International Standard Classification of Occupations (HISCO). The first digit represents major groups classification. Minor groups classification corresponds to the first two digits of the code. Occupations specific to the Egyptian context that do not have an equivalent in HISCO were assigned special codes under a given HISCO major or minor group.</p>
EG1848A_OCCTITLE	EG1848A_OCCTITLE	Occupation title (string)	<p>Person variables Occupation _ _ _ _ _ Respondent's occupation coded using the Historical International Standard Classification of Occupations (HISCO). The first digit represents major groups classification. Minor groups classification corresponds to the first two digits of the code. Occupations specific to the Egyptian context that do not have an equivalent in HISCO were assigned special codes under a given HISCO major or minor group.</p>

ID	Name	Label	Question
EG1848A_ORGPROV3	EG1848A_ORGPROV3	Province of origin	Person variables Province of origin __ The respondent's province of origin, which could coincide with the location currently living in. This information is self-identified. Thus, for example, it could correspond to their place of birth, place where the respondent's parents were born, or where most of the respondent's family members currently reside.
EG1848A_ORGDIST	EG1848A_ORGDIST	District of origin	Person variables District of origin __ The respondent's district of origin, which could coincide with the location currently living in. This information is self-identified. Thus, for example, it could correspond to their place of birth, place where the respondent's parents were born, or where most of the respondent's family members currently reside.
EG1848A_RELATE	EG1848A_RELATE	Relationship to household head	Person variables Relationship to the household head __ Relationship to the household head.
EG1848A_SEX	EG1848A_SEX	Gender	Person variables Gender Respondent's gender. [] Male [] Female
EG1848A_ETHNIC	EG1848A_ETHNIC	Ethnicity	Person variables Ethnicity __ Respondent's ethnicity.
EG1848A_DISAB	EG1848A_DISAB	Disability status	Person variables Disability status (infirmities) _ Respondent's disability status. [] Without disability [] Disabled
EG1848A_REGSTAT	EG1848A_REGSTAT	Census registration status	Person variables Census registration status (enumeration) Denotes whether the individual was enumerated for the census or not. The value 0 ("not enumerated") includes people who may not be living in the household at the time of enumeration, such as emigrants, incarcerated individuals, individuals residing in the military barracks, or individuals protected by their respective consulates. [] Not enumerated [] Enumerated

ID	Name	Label	Question
EG1848A_PERWT	EG1848A_PERWT	Person weight	<p>Person variables Person sampling weight</p> <p>The person sampling weight is the inverse of the sampling probability: it is equal to the (estimated) population size of the province divided by the actual sample size. The sampling probabilities and person sample weights vary across provinces, because of two reasons. First, Cairo and Alexandria were sampled by design at a higher rate (8-12%) than the other provinces (1%). Second, certain provinces are assigned higher person weights in the post-sampling phase to account for the non-enumeration of neighboring or geographically similar provinces. Specifically, the weights include the following adjustments:</p> <p>(1) Rosetta in both 1848 and 1868 receives a higher weight to account for the missing Damietta. (2) Al-Qusayr in 1848 is over-weighted to represent the missing Suez. (3) Al-Arish in 1868 is over-weighted to represent al-Qusayr and Suez. (4) Al-Daqahliya and al-Sharqiya in 1868 are over-weighted to represent the other missing Lower Egypt provinces (al-Qalyubiya, al-Gharbiya, al-Minoufiya, al-Buhayra). (5) Giza, Bani-Souayf, Fayuum, and Asyut in 1868 are over-weighted to represent the other missing Upper Egypt provinces (Minya, Girga, Qena, Aswan).</p>

Total: 32